SEQUENCE LISTING

<110> Bougueleret, Lydie
      Bairoch, Amos
      Niknejad, Anne

<120> Engineered Human Kunitz-Type Protease
   Inhibitor

<130> 54720-8015.US00

<140> US 10/807,204
<141> 2004-03-22

<150> PCT/EP03/01629
<151> 2003-02-18

<150> US 60/358,683
<151> 2002-02-21

<160> 16

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 131
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(131)
<223> eppin-like precursor

<221> SIGNAL
<222> (1)...(25)
<223> predicted by SignalP 2.0

<221> PEPTIDE
<222> (26)...(131)
<223> mature peptide

<221> DOMAIN
<222> (77)...(127)
<223> Kunitz domain predicted by pfscan

<221> DISULFID
<222> (33)...(61)
<223> predicted disulfide bond

<221> DISULFID
<222> (40)...(65)
<223> predicted disulfide bond

<221> DISULFID
<222> (48)...(60)

1

<223> predicted disulfide bond

<221> DISULFID
<222> (54)...(69)
<223> predicted disulfide bond

<221> DISULFID
<222> (77)...(127)
<223> predicted disulfide bond

<221> DISULFID
<222> (86)...(110)
<223> predicted disulfide bond

<221> DISULFID
<222> (102)...(123)
<223> predicted disulfide bond

<400> 1
```
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
  1               5                  10                  15
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
               20                  25                  30
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
               35                  40                  45
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
       50                  55                  60
Cys Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65                  70                  75                      80
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
               85                  90                  95
Lys Lys Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Cys Gln Gly
              100                 105                 110
Asn Asn Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys
              115                 120                 125
Lys Tyr His
      130
```

<210> 2
<211> 106
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(106)
<223> mature form

<221> DOMAIN
<222> (52)...(102)
<223> Kunitz domain predicted by pfscan

<221> DISULFID
<222> (8)...(36)
<223> predicted disulfide bond

<221> DISULFID

<222> (15)...(40)
<223> predicted disulfide bond

<221> DISULFID
<222> (23)...(35)
<223> predicted disulfide bond

<221> DISULFID
<222> (29)...(44)
<223> predicted disulfide bond

<221> DISULFID
<222> (52)...(102)
<223> predicted disulfide bond

<221> DISULFID
<222> (61)...(85)
<223> predicted disulfide bond

<221> DISULFID
<222> (77)...(98)
<223> predicted disulfide bond

<400> 2
Glu Gly Ile Leu Gly Lys Pro Cys Pro Lys Ile Lys Val Glu Cys Glu
 1   5   10   15
Val Glu Glu Ile Asp Gln Cys Thr Lys Pro Arg Asp Cys Pro Glu Asn
             20   25   30
Met Lys Cys Cys Pro Phe Ser Cys Gly Lys Lys Cys Leu Asp Phe Arg
         35   40   45
Lys Asp Ile Cys Ser Met Pro Gln Glu Ala Gly Pro Cys Leu Ala Ser
     50   55   60
Ile Pro His Trp Trp Tyr Asn Lys Lys Thr Lys Ile Cys Ser Glu Phe
65   70   75   80
Ile Tyr Gly Gly Cys Gln Gly Asn Asn Asn Asn Phe Gln Thr Glu Ala
             85   90   95
Ile Cys Leu Val Thr Cys Lys Lys Tyr His
             100   105

<210> 3
<211> 11201
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(11201)
<223> public sequence AL031663.2 (94000..105200)

<221> CDS
<222> (2043)...(2132)

<221> CDS
<222> (3358)...(3486)

<221> CDS
<222> (6120)...(6191)

3

```
<221> CDS
<222> (6193)...(6291)

<221> misc_feature
<222> (6291)...(6293)
<223> potential stop codon

<221> 3'UTR
<222> (6294)...(7236)
<223> partial

<221> polyA_signal
<222> (7231)...(7236)

<400> 3
tttaaagaat catgtaaaac aggcttacaa attttctttt taaaaatatt ttattttatt      60
ttacgttccg ggatacatgt gcaggacctg caggtttgtt acataggtaa gcctgtgcca     120
tggtggtttg ctgcacctat caacccatta cctgggtatt aagccccgca tgcattagct     180
atttatcctg atgctttccc tcccctgcc ccctcaacag tccacgtgtt ctcattgttc     240
atctcccact tgcaagtgag aacatgtggt atttactttt gttcctgcgt tactttgctg     300
aggataacgg cttccagctt catccatgtc cctgcaaagg acgtgatctc atttcttttt     360
atggctgcat agtattccat ggcatatatg taccacattt tctttatcca gcctatcgtt     420
gatggatatt tggtttgatt ccatgttttt gctattgtaa atagtgctgc agtgaacata     480
cacatgcatg tatctttata aaagaatgat ttatattcct ttgagtgtat acccagtaat     540
gggattgctg ggtcaaatgg tatttctggt tctagatctt tcaggaattg gcacactctc     600
ttccacaatg gttgaaccaa tttacattcc caacagtgta aaagcattca tgtttctcca     660
cagctttacc agcatctgtt gtttctggac tttttaataa tcaccattct gactggcatg     720
agatggtatc tcattgtggt tttgatttgc atttctctga tgatcactgg tgttgagctt     780
tttttcatat gtttgttggc cgcataaatg tcttcttttg agaagtgtct gttcatgtcc     840
tttgcccact tttcaatggg gttgtttgtt ttttttcttg taaatttgtt taagttcctt     900
gtagattctg gatatcagac cacaaatttt cttaatctgt ctccagtgat gggaaagcta     960
agctatttgt ccactgattc ccaacgccca ctggatgcca gggatactaa atccctcgct    1020
ctttcaaact gttatacttg caagctatac aagctccctg ggggccagag aaaacccaca    1080
gcagaaaatt agggactgat gctagaggta ggaggctgtc aatgtgcatt ggctattatt    1140
cggccataaa aaggaatgaa cttctcactc accactgatg acatgccttg ctgcacacac    1200
accctgctgc caccctgcag aaatgcttcc attacccaca gtcctttgcc agatggaact    1260
gatgcccagg taactggctc ctcacctcct tcgggcttat accaaggata taaaacttgg    1320
tacagatgct caagcgttaa tgcttcaggg tgtgactttt aatccaatgc tgtagctgtg    1380
acaacaggga aaaacggaga acagagttgg agaagtccca aagtaatgaa agaagttatt    1440
tactgttgca aaatcaaacg acttaaagga taaatgtaac tgatctttac atacgaaaga    1500
agagagaggg aagtgaaggg cggggagaaa gagactcaga cagacagaga ctccctgaga    1560
aaaagacagg cacagagtaa aagacataca gatgcacaga aaagagaaga aacagagtga    1620
gaaaaataaa acataagaat agacatgtca ggagaagagt gaaattagaa agaaggaaga    1680
gagctagcat actgttaacc tgcaaactcc ccctgggaaa tggaagactc tggagttaaa    1740
cctcccccttc tgctaaggga gcctggaaag cccaagacaa catgtgccct cttcccctgg    1800
tagaccagaa agcctcagat tgctgaggat taggagccac ctatccccca gtgtaggtac    1860
caggtagtcc cactgatggc ctcaggggatc tggtgttcct gatcggggag tctggctggc    1920
ttgtctccag ggctgtctct taaggtgctg gccacaccg gggcaggcaa agtgcagatg    1980
taggagctgc tgaggagcag ttcttagttt tggtgccatc aaccaggcca gtacttccta    2040
aa atg gga ctc tca gga ctt ctg cca atc ctg gta cca ttc atc ctt        2087
   Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu
    1               5                  10                  15

ttg ggg gac atc cag gaa cct ggg cac gct gaa ggc atc ctt ggc          2132
Leu Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly
             20                  25                  30
```

4

```
agtaagtatt gggaactccg tccttgcaat ggggaaagga aactcagtgg ggaggaagga    2192
acaaagaata attttgcagc tgagaaatta ttttctccac ctctgcctgg aattccaggt    2252
gggaaagagg ggagctattt gaagtctctt ggtaagagac tcccttggtg aatacaggca    2312
tccaaatctg tggcaagtga ctccttgctg ccattccaaa gagttcttcc tggaggtgcc    2372
tggacctaga gctcatgaaa tgaactaaga aggagtagta agcccaagga agtctcccac    2432
aagaaaaaca ggctgctccc cacaacacca cctgtgtcaa cttccaaatg ttctcattgc    2492
agcccccaca cccataccca cagaggaccc tggtctttgg gttaaagaat ggtctccaga    2552
agtcccttgg aatgctgaaa tgtgaagact cacaattagc agctaactgt actgggcact    2612
gaaaattagt ggttgaaagt gagcactcag gagccacaca ttgtaggttc aaacctggct    2672
ccacccctga gtagctgtgt gatcttgggc aaattcacat aatctcctta tgcttccaat    2732
tcttctgttt acaaaatgtg aatattaaga gtacttaccc atagatttgt aatgaggatt    2792
aaggagaaaa ggcagctcac agtaagtgct tcatccatat tactcattgt tacgttacat    2852
tgccccaggc agtttgacga tggcttgtca aactgtgttg ctggttagtc ttcaaagcaa    2912
ccctatgggc aaatgaggaa atcaaggctc agagagagtg cctggcacac caagcgtctc    2972
acagaattca aaagcccagc tatctgatgc cagaaatgtt aatctcaatc atcgagagcc    3032
acagcacctc ctggggggagc aaccaagcac agcaaggctg agtgaccaga acagtcttag    3092
aggaatgcgt ctttctctca ggaatctcat agatggtgcg ggtagggggt aggggttggt    3152
ttaggcatcc gatggtgagt tgttatgggg ttggggatgt gagcagtgtt gtgaaattcc    3212
tgagccccct gagggccact accttctggg gataggtgaa gtggaaacaa agttctctaa    3272
ctctggaggg gtgccctggt ctcggggagg gggagcactg gctcctttaa ggccttgagg    3332
gaatatatct tcccctccct tagag ccg tgt ccc aaa atc aaa gtg gaa tgc    3384
                            Pro Cys Pro Lys Ile Lys Val Glu Cys
                                                        35

gaa gtg gaa gaa ata gac cag tgt acc aaa ccc aga gat tgc cca gaa    3432
Glu Val Glu Glu Ile Asp Gln Cys Thr Lys Pro Arg Asp Cys Pro Glu
 40                    45              50                    55

aac atg aag tgt tgc ccg ttc agc cgt gga aag aaa tgt tta gac ttc    3480
Asn Met Lys Cys Cys Pro Phe Ser Arg Gly Lys Lys Cys Leu Asp Phe
                    60              65                    70

aga aag gtaactcaga tgcttcctaa attacccagt gccctcacct cctatctcca    3536
Arg Lys


cctgcactgc actacctctg tctcctagtt cactgatggc tggtctctga ccaagagttt    3596
aggtacatct tctacttgac acaaatctgc cagtcctgga cttgggagac atggatttta    3656
tttcatttct tcctctattt tggtaagtga tgtgaacaat gtagtttttcc tttctgggcc    3716
tcagtttccc ttttccataag acaggtgtgg ataatggacc acatggtctc ttggctactc    3776
ttcagtgatt ctcgagatgc caagtagtca ccaaagtctc atcaaatgcc aacattgtgt    3836
caggttatat gatagctata agagacaggt atataatctg tatttcctgg cttttttctc    3896
taggcctctc tgtctacata aatccaggat gccttccctg accaccttag accgcctaca    3956
ctctttcttc tctgcccata ttattgagct ctaaagaaga tcatcaagga ccattcgttc    4016
attctattca ttcaacaaat attgattggt gccctgaca cccatgatgg gttccacttg    4076
ctcattgtgg ccaggaaggt cagaaaagtt tttctgggaa agagatccat gatttaagat    4136
ctgaaagatg agtacttaga ccaaaaatgg gaagaaatag cctttcaggc cagataatag    4196
gatgtgcaaa ggctctgtgg caggaggcag gatgttcagg tctagaaagg ctagaaagca    4256
gagagtgagc aagagacaag aaacaggaaa ctggtagtgt atattcaggg caacttgaat    4316
ctgtattccc aagtggccat cctcaagttt tggggtcaaa taaatgccct acttaatgat    4376
attttctgaa tcttgttact ttaggttgac atacctttca agacaataac taggttttca    4436
gtgctttgat aaaacctcaa atgtgtggcc ttccagccag gcatctgtgg tttcatattg    4496
tttaagcatt tgttgcatcc ctatggattg cagctctcta taggttcgct gttgctgaac    4556
caagaaataa ggcatacaca gatttctaga gaaattgtaa ttctataaca gattttcagg    4616
agatgaggaa ggctggtatc tttcagagta caaagtgatt ttggaattga aagaatttct    4676
ttgggttaaa gtttgtcact gacttgtgtt cctgaaccat gaagcatgaa tatgtgggct    4736
```

5

```
gagaaatagt ttctcttgat aaataaacaa gtaacaaatt ttaaaaagag aggaagatga    4796
agctgggcta gttgggggaa ggagagctgg taagggatct ggtcttgatc ctaacagcaa    4856
tggattaacc tatgttcaaa tcccttacca tctgtgtggc ctcaagaaag ttatttagcc    4916
tctctaaacc ttagttttct catctgcata atgagaacaa tgagaggacc tactttacaa    4976
gattaaatgt taatagatgt caactactaa gaacactaat caatgctata attgttgttt    5036
tgttgtgctg acactaatca atgctacaat tgttgttttg tcgtgcttca gacctgaggt    5096
tctaaagaac tctatttgta cctcatagca ctatgggcaa ataattgct caataaatgc      5156
ttaagatcag ggtagatttg agttttgggt tttttttttt ttctgtgaat actcctgaag   ·5216
catcctctct atgccaaatt acctgctaga aactgctgta tcaagatggc ataagatgca    5276
atcccctctc aaaaacctca cattctcatg gaggaatcag aaaatatgaa caagaaagta    5336
aatggaaatg ataaatgtgt gtgcaagact gagaagcaac gaagaaaagc aaagagtaaa    5396
gaatgaatcc aggaaagact gcataaaaga gtcaatgtcc ggaaagggtt ttaaagaatg    5456
aacaagagtt cacttagcaa gaggggctg gatatgtcag gcatgtgcaa aggcacgggg     5516
tgtgtaagaa agtgctattt ttagagcata agtttgggaa gttgggtgtt ggggcagcat    5576
aagaaataaa gctggagagg agaaggagaa aatggattct gaagtccctc tgggtttggt    5636
agaactttat cctgtgccca gcagatcggc cattctggaa accaatgttg agcagattgg    5696
agaagtcttg agatcagct ggtggccttt catggagtcc ttgcaaaaga taattatgct      5756
tgaattaaga aagcagccgg aaggaaataa aggagagcta tttaggaggc aaaatcagcc    5816
agatttagtg ttcagatgtt ggggaggcaa ggaaaaagaa gccatgtgat gtcacagctg    5876
ctccttccaa cagccctgtt aggcgagcag gttgggattt tgtattctca tcttacagat    5936
gaggaagtca aggaccagag aggtaaaatg agtggtcagg gttcacactg aattgggaac    5996
agagccctga ggcaaatttg cttctcagtg gagacaaagg gaatggtggc cagagaaggg    6056
gagagatgga cattctccag ggagacaggc acccaaacag ctctattctg cgatctccct    6116
gca gat ata tgc agt atg cca cag gag gct ggc ccc tgc ctg gcc tcc       6164
    Asp Ile Cys Ser Met Pro Gln Glu Ala Gly Pro Cys Leu Ala Ser
        75                  80                      85


ata cca cac tgg tgg tac aat aaa aaa a act aag atc tgc tcc gaa ttc     6213
Ile Pro His Trp Trp Tyr Asn Lys Lys   Thr Lys Ile Cys Ser Glu Phe
   90                  95                      100


atc tat ggc ggt agc cag ggg aac aat aac aac ttc caa act gaa gct      6261
Ile Tyr Gly Gly Ser Gln Gly Asn Asn Asn Asn Phe Gln Thr Glu Ala
105                 110                 115                 120


atc tgt ctg gtc acc tgc aaa aaa tac cat aagtcccaga ggtcccggtc         6311
Ile Cys Leu Val Thr Cys Lys Lys Tyr His
                125   ·                130


tcctgtgctc accaaggcca cactgggagg tctgggtgtt ggctggtcta ttccaagacc    6371
tgggtggcgc tggggatgac aaaaccagct ccaatgcaga agtataagta gaaggatatt    6431
ttgggaaaga gggtgggaag ggagggatta gtcaaaggga tattggcaag tatgaggtga    6491
gtagtgggtg tagagagaaa acagaagtgg tggagtatcc agaccaggt cagacggaag      6551
cccggtaaac ccagcccagc cctgggcacc attcatcagc caatcattat agtcctttac    6611
ttctcactaa accttgttgc tacttctctt cctttgaaag gttatttcta accagggcaa    6671
ccacatactt tattggccaa accaaatcac ttttgaaagt gtctcaaggt gaggtgccat    6731
taattattac actgagacaa caggcataaa ctgggactct actggacaag tcagaactca    6791
tgatcattct aggagccccc aaactcacct tcattccatt cctgcccaaa gatgtaaaaa    6851
tgatcccacc tccttttttcc cattaggtgc aagatttggt tcctaatgtg gtcagggtcc    6911
aagcatctca ccctttattc ttctgcttca caggtcagcc ttactttata ccataaggag    6971
gagcttgaat aacctccagg atttggctca taatccaggc ctctctccac gtgtgcctga    7031
ttgatgctcc aaattggctt ccacgggcca aaccttggct gttccagaaa ctgaacccca    7091
ggaattgctt acacactttc ttccagcgta gcatctcttc aaacacaatg ctcttcccct    7151
tgaccacttc tcagtatgaa actctatgtc ttcaactttc agaccccca tttatttgta      7211
tgaaggcttc agttgcctta ataaagtgct ctgtggggtc cgatctctcc gaggctgaat    7271
aattcacatc tgcaatatgc tgatattgac atacactcac atctctctat atgcaatttt    7331
cagagaacaa aaagccattg cattttctgg tgacaccaaa aaagggagag agtatgaaaa    7391


                                     6
```

```
tcctgcaaat tgagaatgct agtgacattg ttactttggt aaagttgtcg ttaagatgct   7451
gaagaatgga gccagtgcag atccagattt tgtatggtct aaacctgata aatgaggagc   7511
tcctttagaa accaagcaca ggctgggcac ggtggctcac gcctgtaatc ccagcacttt   7571
gggaggccaa ggagggcgga tcacctgagg tcaggagttc aagaccagcc tgaccaacgt   7631
ggagaaacac catctctact aaaaatacaa aattagccag gcgtggtggt gcatgcctgt   7691
aatcccaact actcgggagg ctgaggcagg agaattgttt gaacccagga ggcaaaagtt   7751
gcagtgagct gagattgcac cattgcactc cagcctgggc aataagaaca aaactcaaaa   7811
aaaaaaaaaa aaaaaaaga aagaaagaaa gaaagaaaaa aagaaagaaa ccaagcacaa   7871
atgagtacaa aagagaatat tatttcaaat gaggagggaa attgcaacca gctgcaaatt   7931
tttaaaggct gacaaacagt aaaaattgca caaaacccat ataaattaat aattcattgt   7991
ttacatttat gacctcaatt ctgtaaagca ctatttctgt ggagagaatg gaaagataac   8051
tttttatttta atgtggggat caaaatatga tattttaatt tttattatag gtgaattgga   8111
taaaacatgt acctgaatgc ttgaagccat gcttgatgtc tgcagttctg ttacaggttt   8171
gtgtcctaga aacacaggaa ttttaataaa ttatcttttg tacaactcta ttaaaaaaag   8231
aaaaaatgca tggtacattt aaaattgtac atattgaatt atcaagtata tgcttactac   8291
aaaagaactt ctgaacaggc ttccataaga gctgattttt tttttctaat taaaatgtga   8351
tgtatctgac cagactcagt ggttcatgcc tgtaatccca gcactttggg aggccgaagg   8411
gggcagatca cctgatgtca agagttcaag accagactgg ccaacgtggt gaaaccccat   8471
ctctactaaa aatacaagaa attagccagg tgtggtggtg cacacttgta gccccagcta   8531
ctcaggaggc tgaggcagga gaattgcttg aacccgggag gtggaggttg cagtgagcca   8591
agatcgtgcc attgcactct agcctgggcg gtggagtgag actccatctc aaaaaaaaaa   8651
ttgatgtatc tgatatttga aagaatgcta cacaggcttg ttttcctggc tccacacatt   8711
tcaaaacttg tttctcctct accacccgca accttcttgg gctaagaaca ttcattttac   8771
agtgcaacct ctagtcccat aacttagtgt cataacatca agggaatagg cccagtagaa   8831
aacaggagta ttatttgcag ctgcttctcc accaggacag ctagcagcaa cttcctacac   8891
atacaaatga cagctaacga cctaaatata ttctactaac tccaaaataa atatatcccc   8951
aactcaactt cctttttaatt tgcatgccta aaatgctggc agtcgcccaa acaccacctg   9011
acctaggaga agtgtgacag agggaaaggt cagagtggaa agggaaggtg gattggcgat   9071
tatggttaga atatcttact tttgcagata ttacaagtca tctgaccatg taaacatata   9131
gctagaagca ctcccagaat cttaaaaggg gctcgctcga gcgaggaccc tgaggcttgg   9191
ctccatcgac ctcatggtaa gtccgtctct gaaagggctc atggacttgt ttactggagt   9251
tgtagtaaag aagacagttc catagtatgg tgcagaaact ccaaaattga aaatagctgg   9311
acaggcatga aagggggaga tggaggaaag cagagaagag ttaaagagga agggagaaga   9371
aagggtgaa ggaggaaaag acactggatg gtgtgactgt tttaaagcaa ggtcctaaaa   9431
ttctttcact caccttccac ctagaagtgg gatctatgtc ccctcctttt gaatctgaga   9491
aaactctgtg actgctttga ccaagagaat aaagcaaaaa tgatgctatg tgacctccaa   9551
ggctaagcta gaataagtgc aagtcccact tgatcctctt atgtagcttg tttgggggca   9611
gtgggaaact gccatgtaag atgtccaact actctgagag aggcccatgc tagagagacc   9671
catgtgggtg ctgtgatcag caggcccaga taagctctga ctgatagcca gcatcaattt   9731
cagccatgca aatgagatgc ccaacccaat agagccttca gatggctgca gccccagcag   9791
acatctgact gcaaccatgg gaaagaccct ataaagaac caccaagcag attgcttccc   9851
agatggctga cccacaatat aaaatggttg ctttaaagtg ctaagttttc agaaaattta   9911
ttcacagaat aggtgatcaa agatggaaac gggctaggag cagtggctca cagctgtaat   9971
cccaggcact ttgggagtcc aaggcaggca gataacttga agtcagaagt tcaagaccag  10031
cctggccaaa acggcaaaac cctgtctcta ctaaaaatac aaaaattagc cagacatagt  10091
ggcgggcacc tgtaatacca cctactcggg aggctggggc aggagaatct cttcagcccg  10151
ggaggcggag gttgcagtga gctgagattg cgccactgca ctccagagcg agatttggtc  10211
tcaaaaaaaa aaaagaaag aaaaaaaga tgaaaaggga gtagaaggaa ggataggag  10271
agagaaagaa gaaagagaca cccataggac tgggcaagat aagaggatgc caggcttgag  10331
aagggaacct cacaacaggc agcttgtcat cacagcacac caagacttgc tatctcttgt  10391
caaataagtt tgttgtttaa tataacttgg atcaaggggc tctaaagagg ctggaattca  10451
tttgctgttg atgctgttgc catctctgct tccctctctt tagaaagggc tggtaaaact  10511
gattgtagaa aactctaaag gcactgcaaa aaaaggttag aattaataaa caaatttatt  10571
aaagttacag gatacaaaat caacatacaa aaatcaatag caactttata caccaagaac  10631
aatctcattt ccatcagcgg caacaacaac aaaacactca gtaattaact caaccaaaga  10691
agtgaaaaac ttatgcattg acaattataa aacatcgatg aaggaaattt aaaagacaca  10751
aataaatgaa aagacatcca tgctcataga ttagaagaat caatgttgtc aaaatgtcca  10811
```

7

```
cactacccaa  agcaatctac  agagcacatg  caatttctgc  caaaatgcca  atggcatttt  10871
ttacagaaac  ataaaaagca  atgctaaaat  tcatatggaa  ccacaaaaga  ccccaaattg  10931
ccaaaggaat  tttgaaaaag  aacagagctg  taagcaccac  atttcctgat  gtcaaattac  10991
attgcaaagc  tagagtaatc  aaaagagtgt  ggtactggca  taaaaacaga  aaagcagaca  11051
tacaggccaa  tggaacagaa  tagagaagcc  agaactaagc  ccacccattt  atggtcaatt  11111
gatctttgac  gagggtgtca  aaaacacaca  ataggcaaca  gaagatctct  ttaacaaatg  11171
gtgctgggaa  aactagatat  tcacatgcaa                                      11201
```

<210> 4
<211> 30
<212> PRT
<213> Homo sapiens

<400> 4
```
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
 1               5                  10                  15
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly
            20                  25                  30
```

<210> 5
<211> 43
<212> PRT
<213> Homo sapiens

<400> 5
```
Pro Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln
 1               5                  10                  15
Cys Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe
            20                  25                  30
Ser Cys Gly Lys Lys Cys Leu Asp Phe Arg Lys
        35                  40
```

<210> 6
<211> 24
<212> PRT
<213> Homo sapiens

<400> 6
```
Asp Ile Cys Ser Met Pro Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile
 1               5                  10                  15
Pro His Trp Trp Tyr Asn Lys Lys
            20
```

<210> 7
<211> 33
<212> PRT
<213> Homo sapiens

<400> 7
```
Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Ser Gln Gly Asn Asn
 1               5                  10                  15
Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys Lys Tyr
            20                  25                  30
His
```

<210> 8

8
```

```
<211> 1339
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1339)
<223> reconstructed cDNA from SEQ ID NO:3

<221> CDS
<222> (1)...(294)

<221> CDS
<222> (296)...(394)

<221> misc_feature
<222> (394)...(396)
<223> potential stop codon

<221> 3'UTR
<222> (397)...(1339)
<223> partial

<221> polyA_signal
<222> (1334)...(1339)

<400> 8
atg gga ctc tca gga ctt ctg cca atc ctg gta cca ttc atc ctt ttg      48
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
 1               5                  10                  15

ggg gac atc cag gaa cct ggg cac gct gaa ggc atc ctt ggc aag ccg      96
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
                20                  25                  30

tgt ccc aaa atc aaa gtg gaa tgc gaa gtg gaa gaa ata gac cag tgt     144
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
            35                  40                  45

acc aaa ccc aga gat tgc cca gaa aac atg aag tgt tgc ccg ttc agc     192
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
        50                  55                  60

cgt gga aag aaa tgt tta gac ttc aga aag gat ata tgc agt atg cca     240
Arg Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65                  70                  75                  80

cag gag gct ggc ccc tgc ctg gcc tcc ata cca cac tgg tgg tac aat     288
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
                85                  90                  95

aaa aaa a act aag atc tgc tcc gaa ttc atc tat ggc ggt agc cag ggg  337
Lys Lys   Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Ser Gln Gly
             100                 105                 110

aac aat aac aac ttc caa act gaa gct atc tgt ctg gtc acc tgc aaa     385
Asn Asn Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys
```

```
aaa tac cat aagtcccaga ggtcccggtc tcctgtgctc accaaggcca                434
Lys Tyr His
        130


cactgggagg tctgggtgtt ggctggtcta ttccaagacc tgggtggcgc tggggatgac      494
aaaaccagct ccaatgcaga agtataagta gaaggatatt ttgggaaaga gggtgggaag      554
ggagggatta gtcaaaggga tattggcaag tatgaggtga gtagtgggtg tagagagaaa      614
acagaagtgg tggagtatcc cagaccaggt cagacggaag cccggtaaac ccagcccagc      674
cctgggcacc attcatcagc caatcattat agtcctttac ttctcactaa accttgttgc      734
tacttctctt cctttgaaag gttatttcta accagggcaa ccacatactt tattggccaa      794
accaaatcac ttttgaaagt gtctcaaggt gaggtgccat taattattac actgagacaa      854
caggcataaa ctgggactct actggacaag tcagaactca tgatcattct aggagccccc      914
aaactcacct tcattccatt cctgcccaaa gatgtaaaaa tgatcccacc tccttttcc      974
cattaggtgc aagatttggt tcctaatgtg gtcagggtcc aagcatctca ccctttattc      1034
ttctgcttca caggtcagcc ttactttata ccataaggag gagcttgaat aacctccagg      1094
atttggctca taatccaggc ctctctccac gtgtgcctga ttgatgctcc aaattggctt      1154
ccacgggcca aaccttggct gttccagaaa ctgaacccca ggaattgctt acacactttc      1214
ttccagcgta gcatctcttc aaacacaatg ctcttcccct tgaccacttc tcagtatgaa      1274
actctatgtc ttcaactttc agaccccca tttatttgta tgaaggcttc agttgcctta       1334
ataaa                                                                 1339
```

<210> 9
<211> 98
<212> PRT
<213> Homo sapiens

<400> 9

```
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
1               5                   10                  15
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
            20                  25                  30
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
        35                  40                  45
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
    50                  55                  60
Arg Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65                  70                  75                  80
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
                85                  90                  95
Lys Lys
```

<210> 10
<211> 33
<212> PRT
<213> Homo sapiens

<400> 10

```
Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Ser Gln Gly Asn Asn
1               5                   10                  15
Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys Lys Tyr
                20                  25                  30
His
```

<210> 11
<211> 396
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(396)
<223> reconstructed cDNA for SEQ ID NO:1

<221> CDS
<222> (1)...(396)

<400> 11

```
atg gga ctc tca gga ctt ctg cca atc ctg gta cca ttc atc ctt ttg      48
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
1               5                   10                  15

ggg gac atc cag gaa cct ggg cac gct gaa ggc atc ctt ggc aag ccg      96
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
                20                  25                  30

tgt ccc aaa atc aaa gtg gaa tgc gaa gtg gaa gaa ata gac cag tgt      144
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
            35                  40                  45

acc aaa ccc aga gat tgc cca gaa aac atg aag tgt tgc ccg ttc agc      192
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
    50                  55                  60

cgt gga aag aaa tgt tta gac ttc aga aag gat ata tgc agt atg cca      240
Arg Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65                  70                  75                  80

cag gag gct ggc ccc tgc ctg gcc tcc ata cca cac tgg tgg tac aat      288
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
                85                  90                  95

aaa aaa act aag atc tgc tcc gaa ttc atc tat ggc ggt tgc cag ggg      336
Lys Lys Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Cys Gln Gly
                100                 105                 110

aac aat aac aac ttc caa act gaa gct atc tgt ctg gtc acc tgc aaa      384
Asn Asn Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys
            115                 120                 125

aaa tac cat taa                                                       396
Lys Tyr His *
        130
```

<210> 12
<211> 131
<212> PRT
<213> Homo sapiens

<400> 12

11

Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
1   5   10   15
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
        20   25   30
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
        35   40   45
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
    50   55   60
Cys Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65   70   75   80
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
            85   90   95
Lys Lys Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Cys Gln Gly
            100   105   110
Asn Asn Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys
            115   120   125
Lys Tyr His
    130

<210> 13
<211> 133
<212> PRT
<213> Homo sapiens

<400> 13
Met Gly Ser Ser Gly Leu Leu Ser Leu Leu Val Leu Phe Val Leu Leu
1   5   10   15
Ala Asn Val Gln Gly Pro Gly Leu Thr Asp Trp Leu Phe Pro Arg Arg
        20   25   30
Cys Pro Lys Ile Arg Glu Glu Cys Glu Phe Gln Glu Arg Asp Val Cys
        35   40   45
Thr Lys Asp Arg Gln Cys Gln Asp Asn Lys Lys Cys Cys Val Phe Ser
    50   55   60
Cys Gly Lys Lys Cys Leu Asp Leu Lys Gln Asp Val Cys Glu Met Pro
65   70   75   80
Lys Glu Thr Gly Pro Cys Leu Ala Tyr Phe Leu His Trp Trp Tyr Asp
            85   90   95
Lys Lys Asp Asn Thr Cys Ser Met Phe Val Tyr Gly Gly Cys Gln Gly
            100   105   110
Asn Asn Asn Asn Phe Gln Ser Lys Ala Asn Cys Leu Asn Thr Cys Lys
            115   120   125
Asn Lys Arg Phe Pro
    130

<210> 14
<211> 134
<212> PRT
<213> Mus musculus

<400> 14
Met Lys Leu Ser Gly Phe Val Ser Ile Leu Val Leu Phe Gly Leu Leu
1   5   10   15
Ala Arg Val Gln Gly Pro Ser Leu Ala Asp Leu Leu Phe Pro Arg Arg
        20   25   30
Cys Pro Arg Phe Arg Glu Glu Cys Glu His Gln Glu Arg Asp Leu Cys
        35   40   45
Thr Arg Asp Arg Asp Cys Pro Lys Lys Glu Lys Cys Cys Val Phe Asn

12

```
        50                    55                    60
Cys Gly Lys Lys Cys Leu Asn Pro Gln Gln Asp Ile Cys Ser Leu Pro
65                    70                    75                    80
Lys Asp Ser Gly Tyr Cys Met Ala Tyr Phe Arg Arg Trp Trp Phe Asn
                85                    90                    95
Lys Glu Asn Ser Thr Cys Gln Val Phe Ile Tyr Gly Gly Cys Gln Gly
                100                   105                   110
Asn Asn Asn Asn Phe Gln Ser Gln Ser Ile Cys Gln Asn Ala Cys Glu
            115                   120                   125
Lys Lys Ser Ser Leu Thr
        130


<210> 15
<211> 131
<212> PRT
<213> Homo sapiens


<400> 15
Met Gly Leu Ser Gly Leu Leu Pro Ile Leu Val Pro Phe Ile Leu Leu
 1                 5                    10                    15
Gly Asp Ile Gln Glu Pro Gly His Ala Glu Gly Ile Leu Gly Lys Pro
                20                    25                    30
Cys Pro Lys Ile Lys Val Glu Cys Glu Val Glu Glu Ile Asp Gln Cys
            35                    40                    45
Thr Lys Pro Arg Asp Cys Pro Glu Asn Met Lys Cys Cys Pro Phe Ser
        50                    55                    60
Arg Gly Lys Lys Cys Leu Asp Phe Arg Lys Asp Ile Cys Ser Met Pro
65                    70                    75                    80
Gln Glu Ala Gly Pro Cys Leu Ala Ser Ile Pro His Trp Trp Tyr Asn
                85                    90                    95
Lys Lys Thr Lys Ile Cys Ser Glu Phe Ile Tyr Gly Gly Ser Gln Gly
                100                   105                   110
Asn Asn Asn Asn Phe Gln Thr Glu Ala Ile Cys Leu Val Thr Cys Lys
            115                   120                   125
Lys Tyr His
        130


<210> 16
<211> 136
<212> PRT
<213> Mus musculus


<400> 16
Met Arg Leu Trp Gly Leu Leu Pro Phe Leu Val Pro Phe Ile Leu Leu
 1                 5                    10                    15
Trp Ser Ile Gln Glu Pro Glu Leu Ala Glu Gly Phe Phe Ile Arg Thr
                20                    25                    30
Cys Pro Arg Val Arg Val Lys Cys Glu Val Glu Glu Arg Asn Glu Cys
            35                    40                    45
Thr Arg His Arg Gln Cys Pro Asn Lys Lys Arg Cys Cys Leu Phe Ser
        50                    55                    60
Cys Gly Lys Lys Cys Met Asp Leu Arg Gln Asp Val Cys Ser Leu Pro
65                    70                    75                    80
Gln Asp Pro Gly Pro Cys Leu Ala Tyr Leu Pro Arg Trp Trp Tyr Asn
                85                    90                    95
Gln Glu Thr Asp Leu Cys Thr Glu Phe Ile Tyr Gly Gly Cys Gln Gly
                100                   105                   110
```

```
Asn Pro Asn Asn Phe Pro Ser Glu Gly Ile Cys Thr Val Val Cys Lys
        115                 120                 125
Lys Lys Gln Met Ser Ser Trp Ile
    130                 135
```